

PENERAPAN ALGORITMA *K-MEANS* PADA PENGELOMPOKAN DATA SISWA BARU BERDASARKAN JURUSAN DI SMK NEGERI 1 KOTA JAMBI BERBASIS WEB”.

Faiza Rini¹⁾, Novhirtamely Kahar²⁾, ³⁾Juliana

¹⁾ Sistem Informasi STMIK Nurdin Hamzah Jambi

²⁾ ³⁾ Teknik Informatika STMIK Nurdin Hamzah Jambi

Jl. Kolonel Abunjani Sipin Jambi

Email: faiza_rini@yahoo.co.id, novmely@ymail.com, juliana.tandy@yahoo.co.id

Abstrak

Pemilihan Jurusan Bidang Keahlian di Sekolah Menengah Kejuruan(SMK) menjadi bagian penentu keberhasilan siswa dalam menyelesaikan studinya. Sistem yang akan penulis bangun mengikuti perubahan aturan seleksi penerimaan siswa baru, yang sebelumnya menggunakan sistem pembobotan berdasarkan jurusan, menjadi seleksi hasil Nilai akhir pada SKHU dari SMP asal. Langkah yang sangat penting untuk memperoleh hasil maksimal pengelompokan jurusan adalah dengan pengolahan data. Salah satu contoh pengolahan data adalah dengan pengklasteran. Salah satu algoritma pengelompokan data adalah algoritma *K-Means*. Algoritma *K-Means* adalah algoritma *clustering* yang paling sederhana dibanding dengan algoritma yang lain. Algoritma ini mempunyai kelebihan yaitu mudah diterapkan dan dijalankan, relative cepat, mudah untuk diadaptasi, dan paling banyak dipraktekkan dalam data mining. Algoritma ini termasuk salah satu algoritma paling penting dalam data mining.

Aturan dan metode algoritma *K-Means* dapat diterapkan pada sebuah program bantu untuk mengelompokkan data nilai ujian nasional siswa. Adapun sistem yang dirancang adalah penerapan algoritma *K-Means* untuk pengelompokan siswa baru berdasarkan nilai ujian nasional. Langkah-langkah algoritma *K-Means* diterjemahkan kedalam bahasa pemrograman untuk melakukan tugas *clustering* data. Data nilai ujian calon siswa baru yang telah dimasukkan kedalam input program, kemudian program melakukan pengolahan seleksi berdasarkan peringkat dibandingkan kuota global kapasitas penerimaan sekolah, hasil lulus seleksi akan dialgoritma *K-Means*, dan hasilnya berupa *cluster* data nilai. Implementasi sistem yang dirancang dengan menggunakan pemograman Web, basisdata dengan menggunakan *SQLyog*.

Adapun bentuk-bentuk perangkat lunak yang akan diimplementasikan terdiri dari implementasi antarmuka yaitu login, menu utama, tampilan input data pendaftar, input data nilai, input data admin, input data jurusan, input data periode, tampilan proses seleksi, tampilan data tampilan proses *clustering*, tampilan laporan *cluster*, tampilan laporan pendaftar, tampilan laporan lulus seleksi, dan grafik *cluster*.

Keyword : Algoritma *K-Means*, *cluster*, Pemograman Web

1. Pendahuluan

a. Latar Belakang

SMK Negeri 1 Kota Jambi merupakan Salah Satu Sekolah Menengah Kejuruan di Kota Jambi, dan termasuk sekolah terbanyak peminatnya. Sekolah Ini mengapuh bidang Bisnis Manajemen dan Teknologi Informasi. Adapun Jurusan yang terdapat disekolah ini antara lain: Akuntansi, Administrasi Perkantoran, Pemasaran, Serta Teknik Komputer dan Jaringan. Tiap tahunnya SMK Negeri 1 membuka pendaftaran sesuai instruksi pusat dalam jangka waktu 6 hari kerja. Dengan calon peserta yang banyak serta sistem yang

digunakan masih manual, menimbulkan kendala waktu yang tidak efisien, tenaga administrasi yang banyak, serta pengarsipan data yang kurang efektif. Seringkali dijumpai siswa yang hanya memilih jurusan berdasarkan informasi dari teman, tidak memperhatikan nilai dan peminatan kurikulumnya sehingga menjadi salah jurusan dan berakibat putus sekolah di tengah jalan, seperti halnya Pemilihan Jurusan Bidang Keahlian di Sekolah Menengah Kejuruan(SMK). Sistem yang akan penulis bangun mengikuti perubahan aturan seleksi penerimaan siswa baru, yang sebelumnya menggunakan sistem pembobotan berdasarkan jurusan, menjadi seleksi hasil Nilai akhir pada SKHU dari SMP asal. Maka pengolahan data merupakan langkah yang sangat

penting untuk memperoleh hasil maksimal pengelompokan jurusan. Salah satu contoh pengolahan data adalah dengan pengklasteran. Salah satu algoritma pengelompokan data adalah algoritma *K-Means*. Algoritma *K-Means* adalah algoritma *clustering* yang paling sederhana dibanding dengan algoritma yang lain. Algoritma ini mempunyai kelebihan yaitu mudah diterapkan dan dijalankan, relative cepat, mudah untuk diadaptasi, dan paling banyak dipraktekkan dalam tugas data mining. Algoritma ini termasuk salah satu algoritma paling penting dalam data mining. *K-Means* membagi data kemudian mengelompokkannya kedalam beberapa *cluster* yang memiliki kemiripan dan memisahkan setiap *cluster* berdasarkan perbedaan antar masing-masing *cluster*. Algoritma ini telah dikemukakan oleh beberapa peneliti dari disiplin ilmu yang berbeda. Aturan dan metode algoritma *K-Means* dapat diterapkan pada sebuah program bantu untuk mengelompokkan data nilai ujian nasional siswa. Langkah-langkah algoritma *K-Means* diterjemahkan kedalam bahasa pemrograman untuk melakukan tugas *clustering* data. Data nilai ujian calon siswa baru yang telah dimasukkan kedalam input program, kemudian program melakukan pengolahan seleksi berdasarkan peringkat dibandingkan kuota global kapasitas penerimaan sekolah, hasil lulus seleksi akan dialgoritma *K-Means*, dan hasilnya berupa *cluster* data nilai. Hasil berupa *cluster* data nilai inilah yang digunakan sebagai bahan pertimbangan kebijakan pembagian jurusan pada tahun ajaran tersebut. Pendidikan berbasis kompetensi menekankan pada kemampuan yang harus dimiliki oleh calon siswa agar dapat mengikuti kompetensi dan kurikulum sesuai jurusan. Berdasarkan latar belakang yang telah dijelaskan di atas, maka solusi yang bisa digunakan untuk mengelompokkan data nilai calon siswa baru agar fungsi nilai dapat bermanfaat sebagaimana penentu jurusan dan bidang keahlian dengan merancang sebuah program baru yang dapat membantu dalam mengelola data nilai calon siswa

b. Rumusan Masalah

Bagaimana Merancang dan Membangun Penerapan Algoritma *K-Means* untuk Pengelompokan Siswa Baru di SMK Negeri 1 Kota Jambi Berbasis WEB.

c. Tujuan

1. Untuk mengolah dan mempermudah penyeleksian dan pengelompokan penerimaan siswa baru di SMK Negeri 1 Kota Jambi
2. Untuk membantu panitia penerimaan Siswa baru dalam pengelolaan data nilai calon siswa.

2. Pembahasan

Metode Analisis

Analisis kebutuhan mencakup pekerjaan penentuan kebutuhan atau kondisi yang harus dipenuhi dalam suatu sistem baru atau perubahan sistem, yang mempertimbangkan berbagai kebutuhan yang bersinggungan antar berbagai pemangku kepentingan. Kebutuhan dari hasil analisis ini harus dapat dilaksanakan, diukur, diuji, terkait dengan kebutuhan yang teridentifikasi, serta didefinisikan sampai tingkat detail yang memadai untuk desain sistem. Analisis merupakan bagian yang paling penting dalam membangun algoritma *K-Means* untuk pengelompokan siswa baru berdasarkan Nilai ujian nasional agar pembangunan sistem sesuai dengan yang diinginkan dan dapat memberikan hasil yang efektif. Kemudian pada sistem yang akan dibangun ini harus dilakukan agar tidak terjadi kesalahan, karena kesalahan dalam hal analisis ini dapat menyebabkan kesalahan pada tahap-tahap berikutnya yang juga akan berdampak pada pembangunan sistem menjadi tidak sesuai dengan keinginan.

Adapun penerapan Algoritma *K-Means* untuk siswa baru berdasarkan Nilai ujian nasional ini akan dibangun dengan menggunakan bahasa pemrograman Web, Php. Metode yang digunakan dalam analisis kebutuhan sistem ini adalah metode analisis pendekatan terstruktur. Dengan metode analisis ini maka pembangunan sistem ini akan menjadi lebih terstruktur. Kemudian dengan metode ini modul-modul akan lebih mudah dicoba secara terpisah dan kemudian percobaan dapat dilakukan pada integrasi semua modul untuk meyakinkan bahwa interaksi antar modul dapat berfungsi sebagaimana mestinya. Metode pembangunan sistem yang akan digunakan dalam analisis sistem ini adalah metode yang menekankan pada karakteristik dan data yang akan diproses.

Hasil Analisis

Hasil analisis kebutuhan merupakan pemilihan kebutuhan sistem yang harus diwujudkan dalam perangkat lunak yang meliputi fungsi-fungsi yang dibutuhkan, kinerja yang harus dipenuhi, proses masukan dan keluaran data.

Kebutuhan Masukan

Adapun kebutuhan masukan pada sistem ini adalah :

1. Data Pendaftaran Siswa Baru di SMK Negeri 1 Kota Jambi tahun pelajaran 2015 / 2016.

2. Data Nilai dari Surat Keterangan Hasil Kelulusan (SKHU) sementara yang mencantumkan nilai hasil Ujian Nasional saat SMP.
3. Kuota dari SMK Negeri 1 sebagai bahan pembatas penerimaan dan penyeleksian siswa baru.
4. Data Periode Tahun Pelajaran, Kuota Penerimaan siswa baru.
5. Jumlah *Cluster*, yaitu jumlah *cluster* yang diinputkan pada sistem yang dibuat ini *cluster* dibatasi dengan 4 *cluster* berdasarkan jurusan yang tersedia di SMK Negeri 1 Kota Jambi.

Kebutuhan Proses

Adapun proses-proses yang dibutuhkan adalah sebagai berikut :

1. Proses pengolahan data input.
2. Proses penyeleksian data berdasarkan kuota.
3. Proses pengelompokan data.
4. Proses pembuatan laporan.

Kebutuhan Keluaran

1. Laporan Pendaftar
2. Laporan siswa lulus seleksi
3. Laporan Data Pengelompokan (*cluster*) siswa baru
4. Laporan Grafik Pengelompokan (*cluster*)

Metode perancangan yang digunakan dalam Penerapan Algoritma *K-Means* untuk pengelompokan siswa baru berdasarkan Nilai ujian nasional adalah metode perancangan terstruktur (*Structured Design Method*) dengan menggunakan diagram arus data (*Data Flow Diagram*). Metode Perancangan terstruktur mengarah kepada rancangan yang dapat menghasilkan sistem informasi yang baik. Tujuan utama suatu metode perancangan terstruktur adalah untuk menghasilkan sistem informasi yang menggunakan prosedur dan dokumentasi yang baku dan jelas.

Hasil Perancangan

Hasil perancangan dalam membangun Algoritma *K-Means* untuk pengelompokan siswa baru berdasarkan Nilai ujian nasional merupakan terjemahan perangkat lunak yang dibutuhkan di dalam perancangan, yaitu meliputi perancangan diagram alir data, perancangan *data flow diagram*, basis data dan perancangan antarmuka (*interface*) yang akan dihasilkan.

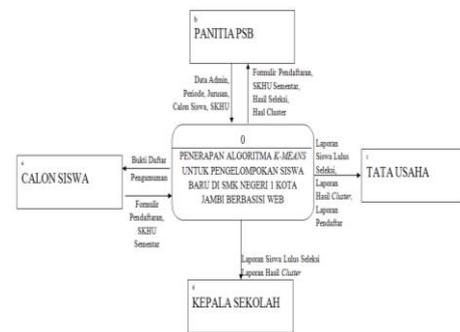
Data Flow Diagram (DFD)

Data Flow Diagram atau disebut juga Diagram Alir Data merupakan penggambaran suatu sistem yang telah ada atau sistem baru yang akan

dikembangkan secara logika yang menjelaskan tentang alur dari suatu masalah yang terjadi pada suatu objek atau suatu masalah yang akan diselesaikan. Pada perancangan ini *Data Flow Diagram* terdiri dari beberapa bagian sebagai berikut :

Diagram Konteks

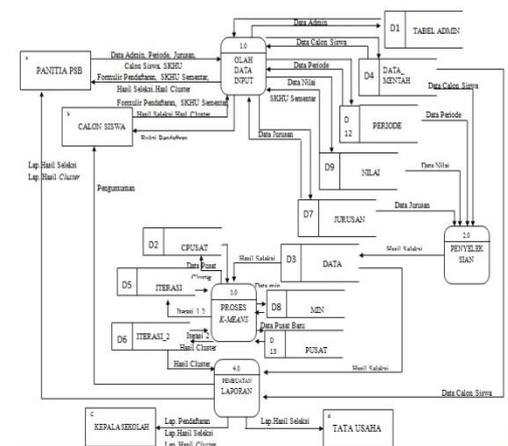
Diagram konteks adalah diagram yang menggambarkan aliran data secara umum, dimana yang ditampilkan adalah proses dan lingkungan luar yang berhubungan dengan proses pengolahan data tersebut, adapun bentuk konteks diagram penerapan Algoritma *K-Means* untuk pengelompokan siswa baru berdasarkan Nilai ujian nasional adalah sebagai berikut :



Gambar 1. Context Diagram

Diagram Level Nol

Diagram ini dibuat untuk menggambarkan tahapan proses yang ada didalam *diagram konteks*, dengan penjabaran yang lebih terperinci.



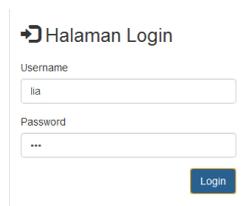
Gambar 2 Data Flow Diagram Level 0

Implementasi sistem yang dirancang menggunakan pemrograman Web, basisdata dengan

menggunakan *SQLyog*. Adapun sistem yang dirancang adalah penerapan algoritma *K-Means* untuk pengelompokan siswa baru berdasarkan Nilai ujian nasional. Adapun bentuk-bentuk perangkat lunak yang akan diimplementasikan terdiri dari implementasi antarmuka yaitu login, menu utama, tampilan input data pendaftar, input data nilai, input data admin, input data jurusan, input data periode, tampilan proses seleksi, tampilan data tampilan proses *clustering*, tampilan laporan *cluster*, tampilan laporan pendaftar, tampilan laporan lulus seleksi, dan grafik *cluster*.

1. Tampilan Antarmuka Login

Pada *interface* Login ini terdapat dua inputan data, yaitu input username dan input password. Dengan menginputkan Username dan Password yang benar maka pengguna dapat mengakses sistem ini.



Gambar 3. Halaman Login

2. Tampilan Hasil Seleksi Siswa

Merupakan bagian dari menu sub Proses Seleksi Menampilkan Hasil Proses Penyeleksian Siswa Baru berdasarkan Periode dan diperingkat berdasarkan kuotanya. Sampel Periode 2015 / 2016 Kuota 320 siswa. Hasil Seleksi ini akan di *cluster* di tahap berikutnya.

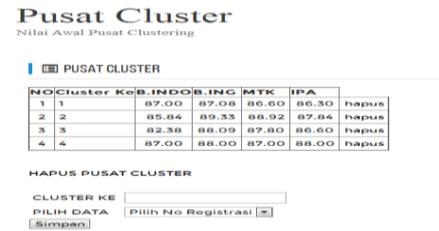
TABEL NILAI SISWA TELAH DI SELEKSI

Kode Jurusan	Nama Jurusan	Bahasa Indonesia	Bahasa Inggris	Matematika	IPA
1	AK ILMAN ANDESTA UTAMA PUTRA	85.30	95.20	95.20	82.00
2	AK ELSA SAVIRA FIDYALOKA	85.30	95.20	90.85	82.00
3	AK NISA LUSIANA HARIA	92.20	93.50	86.24	81.00
4	TKJ MAULANA RIZKY PUTRA SAURI	88.00	90.00	95.00	94.00
5	TN M.ANDRI MAULANA	90.80	92.30	87.00	55.00
6	AP FEBRYAND NURUL FADILLAH	88.50	90.80	92.30	88.70
7	AK HATTA NUGRAHA	88.50	90.80	92.30	88.70
8	AK CINTYA PEBRITA SILABAN	85.30	95.20	86.00	87.46
9	AK NATHANIA CLARESTA ORVILLE	88.00	90.00	93.42	87.54
10	AP FENIA APRILIYA	86.00	90.85	93.50	85.10
11	AK SINTA PERMATA SARI	86.60	90.90	92.20	93.50
12	AK MAI SAPRIADI	90.00	90.00	90.40	89.00
13	AK IRMA YANI	82.10	92.20	93.50	85.10

Gambar 4. Tampilan Hasil Seleksi Data Siswa Awal

3. Tampilan Tampilan Pilih Pusat Cluster

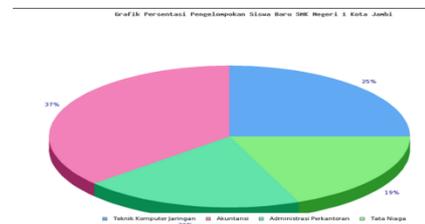
Merupakan bagian dari menu sub Proses Seleksi Cluster, Melakukan pemilihan data pusat yang akan di jadikan centroid, untuk melakukan proses *clustering*.



Gambar 5. Tampilan Tampilan Pilih Pusat Cluster

4. Tampilan Hasil Grafik Cluster Scater Plot

Grafik Lingkaran hasil perhitungan *cluster* Sebagai berikut :



Gambar 6. Tampilan grafik lingkaran

5. Tampilan Laporan Pengelompokan (Clustering)

Laporan yang disediakan pada sistem ini ada tiga salah satunya laporan pengelompokan siswa baru adapun tampilannya sebagai berikut :

LAPORAN CLUSTER SISWA SMKN 1 JAMBI
TAHUN AJARAN 2015/2016
PENGELOMPOKAN KE 1
TEKNIK KOMPUTER JARINGAN

NO	ID Registrasi	Nama Siswa	NIK	JK	B.INDO	B.ING	MTK	IPA
1	PSS-0109	NURFITRA FADLANI	000702148	P	88.30	87.50	90.00	79.07
2	PSS-0116	IVAN NATALIA	011030808	P	85.50	90.00	88.00	78.10
3	PSS-0120	CHARIN WILFA	888600211	P	85.50	90.00	89.00	78.00
4	PSS-0108	HALIM PRISTOWULU	000820084	P	88.30	87.50	90.00	79.07
5	PSS-0109	RENAN EN HARDOYANTIR	000820084	P	85.50	90.00	88.00	78.10
6	PSS-0008	RENAN SYAHPUTRI	000709048	P	85.50	90.00	89.00	78.00
7	PSS-0041	RTS NULUL CAHYA ULAPPAH	000884117	P	85.40	90.00	88.00	78.00
8	PSS-0006	PRIMA SEPTINA HULIMAHAT	000791716	P	86.70	87.00	89.00	78.24
9	PSS-0042	REZZA FARULLI	000844808	P	85.00	90.00	87.00	78.00
10	PSS-0007	RIANCA HILLAH	000744807	L	90.00	92.30	87.00	85.00
11	PSS-0007	IRVINDA HINDA	888881948	P	80.00	92.30	94.00	78.00
12	PSS-0008	NISALUSIANA HARIA	888198703	P	82.00	93.50	93.00	81.00

Jambi,

Drs. R. A. S. I. N. O.
NIP.196606091984421002

Sarjana Magister Sistem Informasi, lulus tahun 2008. Saat ini menjadi Dosen di STMIK Nurdin Hamzah Jambi.

Gambar 7. Tampilan Laporan Pengelompokan
(*Clustering*)

3. Kesimpulan

Setelah melalui banyak tahapan dalam pembuatan sistem perangkat lunak dimulai dari analisis kebutuhan sampai dengan implementasi dan pengujian maka ada beberapa kesimpulan yang didapat tulisan ini yaitu:

1. Penentuan titik pusat (*centroid*) pada tahap awal *Algoritma K-Means* sangat berpengaruh pada hasil *cluster* seperti pada hasil pengujian yang dilakukan dengan menggunakan 320 data dengan *centroid* yang berbeda menghasilkan hasil *cluster* yang berbeda-beda.
2. Semakin banyak data yang digunakan dalam sistem maka semakin optimal *cluster* yang diperoleh.
3. Penerapan metode *Clustering* telah memberikan fungsi yang bermanfaat pada pengelompokan siswa baru, dan dapat menjadi salah satu solusi atau referensi pembagian jurusan pada siswa baru.

Novhirtamely Kahar memperoleh gelar Sarjana Teknik (ST), Jurusan Teknik Informatika UII Yogyakarta. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana Magister Teknologi Informasi UPI "YPTK" Padang, lulus tahun 2015. Saat ini menjadi Dosen di STMIK Nurdin Hamzah Jambi

Juliana Mahasiswi Teknik Informatika STMIK Nurdin Hamzah

Daftar Pustaka

- [1] Badan Standar Nasional Pendidikan. "*Prosedur Operasional Standar Penyelenggaraan Ujian Nasional*": Jakarta : Direktorat Pendidikan, 2016.
- [2] Herawati, Fajar Astuti. "*Data Mining*", Yogyakarta : Penerbit Andi, 2013.
- [3] Kadir, Abdul. "*Buku Pintar Programmer Pemula PHP*": Yogyakarta : MediaKom, 2013.
- [4] Kusriani "*Konsep dan Aplikasi Sistem Pendukung Keputusan*". Yogyakarta : Andi Offset, 2007.
- [5] Prasetyo, Eko. "*Data Mining Konsep dan Aplikasi menggunakan MATLAB*". Yogyakarta: Penerbit Andi, 2012.
- [6] Siberio, Alexander F.K. "*WEB Programming Power Pack*": Yogyakarta : MediaKom, 2013.

Biodata Penulis

Faiza Rini memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Komputer UPI "YPTK" Padang, lulus tahun 2001. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca